

# Integer-Valued APARCH Processes

Maria da Conceição Costa, Manuel G. Scotto, and Isabel Pereira

Departamento de Matemática and CIDMA  
University of Aveiro, 3810-193, Aveiro, Portugal  
{lopescosta,mscotto,isabel.pereira}@ua.pt

**Abstract.** The Asymmetric Power Arch representation for the volatility was introduced by Ding et al. [4] in order to account for asymmetric responses in the volatility in the analysis of continuous-valued financial time series like, for instance, the log-return series of foreign exchange rates, stock indices or share prices. As reported by Brännäs and Quoreishi [1], asymmetric responses in volatility are also observed in time series of counts such as the number of intra-day transactions in stocks. In this work, an asymmetric power autoregressive conditional Poisson model is introduced for the analysis of time series of counts exhibiting asymmetric overdispersion. Basic probabilistic and statistical properties are summarized and parameter estimation is discussed. A simulation study is presented to illustrate the proposed model. Finally, an empirical application to a set of data concerning the daily number of stock transactions is also presented to attest for its practical applicability in data analysis.

**Keywords:** Asymmetric Volatility, Ergodicity, Heteroscedasticity, Overdispersion, Non Linear Time Series, Stationarity

## 1 Introduction

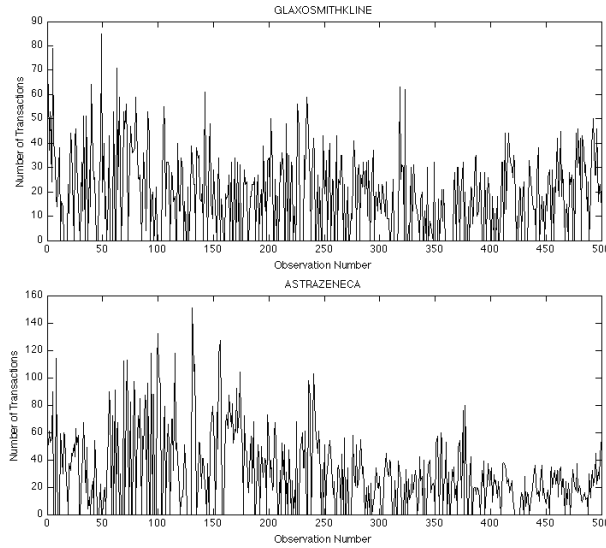
The analysis of continuous-valued financial time series like log-return series of foreign exchange rates, stock indices or share prices, has revealed some common features: sample means not significantly different from zero, sample variances of the order  $10^{-4}$  or smaller and sample distributions roughly symmetric in its center, sharply peaked around zero but with a tendency to negative asymmetry. In particular, it has usually been found that the conditional volatility of stocks responds asymmetrically to positive versus negative shocks: volatility tends to rise higher in response to negative shocks as opposed to positive shocks, which is known as the leverage effect. To account for asymmetric responses in the volatility, [4] introduced the Asymmetric Power ARCH, APARCH( $p, q$ ), in which

$$Y_t = \sigma_t Z_t, \quad \sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \quad t \in \mathbb{Z}, \quad (1)$$

where  $(Z_t)$  is an i.i.d. sequence with zero mean,  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ ,  $\delta \geq 0$ ,  $-1 < \gamma_i < 1$ . The APARCH representation in (1) has some noteworthy

advantages, namely the power of the returns for which the predictable structure in the volatility is the strongest, is determined by the data and the model allows the detection of asymmetric responses of the volatility for positive or negative shocks. If  $\gamma_i > 0$  the leverage effect arises.

Asymmetric responses on the volatility are also commonly observed in the analysis of time series representing the number of intra-day transactions in stocks, in which the numbers are typically quite small as reported in [1]. As an illustration of this kind of data we present in Fig. 1 two time series of count data generated from stock transactions, namely tick-by-tick data for Glaxosmithkline and Astrazeneca downloaded from [www.dukascopy.com](http://www.dukascopy.com). Data consist on the number of transactions per minute during one trading day (19/09/2012, for Glaxosmithkline and 21/09/2012, for Astrazeneca), corresponding to 501 observations for each series. Counts are typically small and both time series contain a large quantity of zeros. After download, data was filtered by the authors in order to fill in the zero counts during the trading periods considered and delete all trading during the first and the last five minutes of each day as trading mechanisms may be different during the opening and closing of the stock exchange market. The existence of this kind of data motivated our proposal for



**Fig. 1.** Time series plots for Glaxosmithkline (*top*) and Astrazeneca (*bottom*).

a counterpart of the APARCH representation of the volatility. To this extent, in Section 2 an INGARCH-type model suitable to account for time series of counts exhibiting asymmetric overdispersion is introduced. Parameter estima-

tion is covered in Section 3. In Section 4 a simulation study is carried out to illustrate INAPARCH(1, 1) model. A real-data application is given in Section 5.

## 2 Integer-Valued APARCH( $p, q$ ) Processes

In this work, focus is put on models in which the count variable is assumed to be Poisson distributed, conditioned on the past, which is to say that the conditional distribution of the count variable, given the past, is assumed to be Poisson with time-varying mean  $\lambda_t$ , satisfying some autoregressive mechanism. An important family of such observation-driven models that is able to handle overdispersion is the class of Autoregressive Conditional Poisson (ACP), first introduced in [10], but also referred to as the INGARCH model due to its analogy to the conventional GARCH model (see Ferland et al. [6]).

An INteger-valued GARCH process of orders  $(p, q)$ , INGARCH( $p, q$ ) in short, is defined to be an integer-valued process  $(Y_t)$  such that, conditioned on the past experience,  $Y_t$  is Poisson distributed with mean  $\lambda_t$ , and  $\lambda_t$  is obtained recursively from the past values of the observable process  $(Y_t)$  and  $(\lambda_t)$  itself, that is

$$Y_t | \mathcal{F}_{t-1} \sim Po(\lambda_t), \quad \lambda_t = \gamma_0 + \sum_{i=1}^p \gamma_i Y_{t-i} + \sum_{j=1}^q \delta_j \lambda_{t-j}, \quad t \in \mathbb{Z},$$

where  $\mathcal{F}_{t-1} := \sigma(Y_s, s \leq t-1)$ ,  $\gamma_0 > 0$ ,  $\gamma_i \geq 0$ , and  $\delta_j > 0$ . In [6] it was shown that the process  $(Y_t)$  is strictly stationary with finite first and second order moments provided that  $\sum_{i=1}^p \gamma_i + \sum_{j=1}^q \delta_j < 1$ . The particular case  $p = q = 1$  was analyzed by Fokianos and Tjøstheim [7] and Fokianos et al. [8] under the designation of Poisson Autoregression. The authors considered linear and nonlinear models for  $\lambda_t$ . For the linear model case, the representation considered is as follows

$$Y_t | \mathcal{F}_{t-1}^{Y, \lambda} \sim Po(\lambda_t), \quad \lambda_t = d + a\lambda_{t-1} + bY_{t-1}, \quad t \in \mathbb{N}, \quad (2)$$

where it is assumed that the parameters  $d, a, b$  are positive, and  $\lambda_0$  and  $Y_0$  are fixed. This representation corresponds exactly to the INGARCH(1, 1) model in [6], nevertheless, the approach followed by [8] is slightly different in the sense that the linear model is rephrased as  $Y_t = N_t(\lambda_t)$ ,  $t \in \mathbb{N}$  with  $\lambda_t$  defined as in (2), and  $\lambda_0$  and  $Y_0$  fixed. For each time point  $t$ , the authors introduced a Poisson process of unit intensity,  $N_t(\cdot)$ , so that  $N_t(\lambda_t)$  represents the number of such events in the time interval  $[0, \lambda_t]$ . Following this rephrasing a perturbation is introduced in order to demonstrate  $\phi$ -irreducibility and as a consequence geometric ergodicity follows. The nonlinear case is considered a generalization of the previous situation in which the conditional mean,  $E[Y_t | \mathcal{F}_{t-1}^{Y, \lambda}] = \lambda_t$ , is a nonlinear function of both the past values of  $\lambda_t$  and the past values of the observations. Sufficient conditions to prove geometric ergodicity were also derived in [8].

It is worth to mention that the models above can not cope with the presence of asymmetric overdispersion. This paper aims to give a contribution towards this direction with the introduction of the INteger-valued APARCH process.

**Definition 1 (INAPARCH( $p, q$ ) model).** An *INteger-valued APARCH*( $p, q$ ) is defined to be an integer-valued process  $(Y_t)$ , such that, conditioned on the past, the distribution of  $Y_t$  is Poisson with mean value  $\lambda_t$  satisfying the recursive equation

$$\lambda_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|Y_{t-i} - \lambda_{t-i}| - \gamma_i (Y_{t-i} - \lambda_{t-i}))^\delta + \sum_{j=1}^q \beta_j \lambda_{t-j}^\delta, \quad t \in \mathbb{Z}$$

with  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$ ,  $|\gamma_i| < 1$  and  $\delta \geq 0$ .

Following the work of Doukhan et al. [5] (see also [2], [12], [9]) we will establish the existence and uniqueness of a stationary solution, and ergodicity for the  $p = q = 1$  case. The INAPARCH(1, 1) process is defined as an integer-valued process  $(Y_t)$  such that

$$Y_t | \mathcal{F}_{t-1} \sim Po(\lambda_t), \quad \lambda_t^\delta = \omega + \alpha (|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta + \beta \lambda_{t-1}^\delta, \quad (3)$$

with  $t \in \mathbb{Z}$ ,  $\alpha \equiv \alpha_1$ ,  $\beta \equiv \beta_1$  and  $\gamma \equiv \gamma_1$ . The  $\gamma$  parameter should reflect the leverage effect relative to the conditional mean of the process  $(Y_t)$ .

**Proposition 1.** *Under the conditions in Definition 1, the bivariate process  $(Y_t, \lambda_t)$  has a stationarity solution.*

*Proof.* For a general Markov chain and according to Theorem 12.0.1(i) in [11], if  $(X_t)$  is a weak Feller chain and if for any  $\epsilon > 0$  there exists a compact set  $C \subset X$  such that  $P(x, C^c) < \epsilon, \forall x \in X$ , then  $(X_t)$  is bounded in probability and thus there exists at least one stationary distribution for the chain. We will show that the chain is bounded in probability and therefore admits at least one stationary distribution. First note that the chain is weak Feller (cf., [2]). Define  $C := [-c, c]$  then

$$\begin{aligned} P(\lambda, C^c) &= P(\lambda_t^\delta \in C^c | \lambda_{t-1} = \lambda) \\ &= P(|\omega + \alpha (|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta + \beta \lambda_{t-1}^\delta| > c | \lambda_{t-1} = \lambda) \end{aligned}$$

which, by Markov's inequality

$$\begin{aligned} &\leq \frac{E[|\omega + \alpha (|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta + \beta \lambda_{t-1}^\delta| | \lambda_{t-1} = \lambda]}{c} \\ &= \frac{E[|\omega|] + E[|\alpha (|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta| | \lambda_{t-1} = \lambda] + E[|\beta \lambda_{t-1}^\delta|]}{c} \end{aligned}$$

and since  $\alpha, \beta, \delta, \lambda > 0$

$$P(\lambda, C^c) \leq \frac{\omega}{c} + \frac{\alpha}{c} E[|(|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta| | \lambda_{t-1} = \lambda] + \frac{\beta \lambda^\delta}{c}.$$

In view of the fact that  $|\gamma| < 1$  and  $|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}) \geq 0$ , the expression above simplifies to

$$P(\lambda, C^c) \leq \frac{\omega + \beta \lambda^\delta}{c} + \frac{\alpha}{c} E[(|Y_{t-1} - \lambda_{t-1}| - \gamma (Y_{t-1} - \lambda_{t-1}))^\delta | \lambda_{t-1} = \lambda].$$

Since by definition

$$\begin{aligned} E [ (|Y_{t-1} - \lambda_{t-1}| - \gamma(Y_{t-1} - \lambda_{t-1}))^\delta | \lambda_{t-1} = \lambda ] &= \\ &= \sum_{y_{t-1}=0}^{+\infty} (|y_{t-1} - \lambda| - \gamma(y_{t-1} - \lambda))^\delta \frac{e^{-\lambda} \lambda^{y_{t-1}}}{(y_{t-1})!}, \end{aligned}$$

then

$$P(\lambda, C^c) \leq \frac{\omega + \beta \lambda^\delta}{c} + \frac{\alpha}{c} e^{-\lambda} \sum_{y_{t-1}=0}^{+\infty} \frac{\lambda^{y_{t-1}}}{(y_{t-1})!} (|y_{t-1} - \lambda| - \gamma(y_{t-1} - \lambda))^\delta.$$

By d'Alembert's criterion, the series  $\sum_{y_{t-1}=0}^{+\infty} \frac{\lambda^{y_{t-1}}}{(y_{t-1})!} (|y_{t-1} - \lambda| - \gamma(y_{t-1} - \lambda))^\delta$ , is absolutely convergent. Being convergent, the series has a finite sum and so it can be written that

$$P(\lambda, C^c) \leq \frac{\omega + \beta \lambda^\delta}{c} + \frac{\alpha}{c} e^{-\lambda} \sum_{y_{t-1}=0}^{+\infty} \frac{\lambda^{y_{t-1}}}{(y_{t-1})!} (|y_{t-1} - \lambda| - \gamma(y_{t-1} - \lambda))^\delta < \epsilon.$$

Thus, for any  $\epsilon > 0$  just choose  $c$  large enough so that

$$\frac{1}{c} \left( \omega + \beta \lambda^\delta + \alpha e^{-\lambda} \sum_{y_{t-1}=0}^{+\infty} \frac{\lambda^{y_{t-1}}}{(y_{t-1})!} (|y_{t-1} - \lambda| - \gamma(y_{t-1} - \lambda))^\delta \right) < \epsilon,$$

leading to conclude that the series has at least one stationary solution.

In proving uniqueness we proceed as follows: first note that the INAPARCH(1, 1) model belongs to the class of observation-driven Poisson count processes considered in Neumann [12],  $Y_t | \mathcal{F}_{t-1}^{Y, \lambda} \sim Po(\lambda_t); \lambda_t = f(\lambda_{t-1}, Y_{t-1}), t \in \mathbb{N}$  with

$$f(\lambda_{t-1}, Y_{t-1}) = (\omega + \alpha(|Y_{t-1} - \lambda_{t-1}| - \gamma(Y_{t-1} - \lambda_{t-1}))^\delta + \beta \lambda_{t-1}^\delta)^{\frac{1}{\delta}}.$$

Thus, the result follows if the function  $f$  above satisfies the following contractive condition

$$|f(\lambda, y) - f(\lambda', y')| \leq k_1 |\lambda - \lambda'| + k_2 |y - y'| \quad \forall \lambda, \lambda' \geq 0, \forall y, y' \in \mathbb{N}_0, \quad (4)$$

where  $k_1$  and  $k_2$  are nonnegative constants such that  $k := k_1 + k_2 < 1$ . For the INAPARCH(1, 1) model the contractive condition simplifies to

$$|f(\lambda_{t-1}, Y_{t-1}) - f(\lambda'_{t-1}, Y'_{t-1})| \leq \left\| \frac{\partial f}{\partial \lambda_{t-1}} \right\|_\infty |\lambda_{t-1} - \lambda'_{t-1}| + \left\| \frac{\partial f}{\partial Y_{t-1}} \right\|_\infty |Y_{t-1} - Y'_{t-1}|,$$

where for the Euclidean space  $\mathbb{R}^d$  and  $h : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $\|h\|_\infty$  is defined by  $\|h\|_\infty = \sup_{x \in \mathbb{R}^d} |h(x)|$ . For the sake of brevity we will skip the theoretical details and conclude that, in the INAPARCH(1, 1) case, if

$$\alpha 2^{\delta+1} \delta + \beta 2^{\delta-1} < 1, \quad (5)$$

for  $\delta \geq 2$ , then the contractive condition holds. This concludes the proof.

Neumann [12] proved that the contractive condition in (4) is, indeed, sufficient to ensure uniqueness of the stationary distribution and ergodicity of  $(Y_t, \lambda_t)$ . The results are quoted below.

**Proposition 2.** *Suppose that the bivariate process  $(Y_t, \lambda_t)$  satisfies (3) and (5) for  $\delta \geq 2$ . Then the stationary distribution is unique and  $E[\lambda_1] < \infty$ .*

**Proposition 3.** *Suppose that the bivariate process  $(Y_t, \lambda_t)$  is in its stationarity regime and satisfies (3) and (5) for  $\delta \geq 2$ . Then the bivariate process  $(Y_t, \lambda_t)$  is ergodic and  $E[\lambda_1^2] < \infty$ .*

Furthermore, following Theorem 2.1. in [5], it can be shown that if the process  $(Y_t, \lambda_t)$  satisfies (3) and (5) for  $\delta \geq 2$ , then there exists a solution of (3) which is a  $\tau$ -weakly dependent strictly stationary process with finite moments up to any positive order and is ergodic.

### 3 Parameter Estimation

In this section, we consider the estimation of the parameters of the INAPARCH(1, 1) model. The conditional maximum likelihood (CML) method can be applied in a very straightforward manner. Note that by the fact that the conditional distribution is Poisson the conditional likelihood function, given the starting value  $\lambda_0$  and the observations  $y_1, \dots, y_n$ , takes the form

$$L(\theta) := \prod_{t=1}^n \frac{e^{-\lambda_t(\theta)} \lambda_t^{y_t}(\theta)}{y_t!} \quad (6)$$

with  $\theta := (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \gamma_1, \dots, \gamma_p, \delta) \equiv (\theta_1, \theta_2, \dots, \theta_{2p+q+2})$ , the unknown parameter vector. The log-likelihood function is given by

$$\ln(L(\theta)) = \sum_{t=1}^n [y_t \ln(\lambda_t) - \lambda_t - \ln(y_t!)] = \sum_{t=1}^n \ell_t(\theta). \quad (7)$$

The score function is the vector defined by

$$S_n(\theta) := \frac{\partial \ln(L(\theta))}{\partial \theta} = \sum_{t=1}^n \frac{\partial \ell_t(\theta)}{\partial \theta}. \quad (8)$$

For the calculation of the first order derivatives of the general INAPARCH( $p, q$ ) model the auxiliary calculations presented below are needed.

$$\frac{\partial \ell_t}{\partial \theta_i} = \frac{\partial \lambda_t}{\partial \theta_i} \left( \frac{y_t}{\lambda_t} - 1 \right), \quad i = 1, \dots, 2 + 2p + q,$$

where  $\frac{\partial \lambda_t}{\partial \theta_i} = \frac{\lambda_t}{\delta \lambda_t^\delta} \frac{\partial (\lambda_t^\delta)}{\partial \theta_i}$ ,  $i = 1, \dots, 2 + 2p + q$ . Thus, for  $i = 1, \dots, p$  and for  $j = 1, \dots, q$ , the first derivatives are given by the following expressions

$$\frac{\partial \lambda_t}{\partial \omega} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta \sum_{i=1}^p \alpha_i g_{t-i}^{\delta-1} (I_{t-i} + \gamma_i) \frac{\partial \lambda_{t-i}}{\partial \omega} + \sum_{j=1}^q \beta_j \frac{\partial \lambda_{t-j}^\delta}{\partial \omega} + 1 \right)$$

$$\begin{aligned}
\frac{\partial \lambda_t}{\partial \alpha_i} &= \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta \sum_{k=1}^p \alpha_k g_{t-k}^{\delta-1} (I_{t-k} + \gamma_k) \frac{\partial \lambda_{t-k}}{\partial \alpha_i} + \sum_{j=1}^q \beta_j \frac{\partial \lambda_{t-j}^\delta}{\partial \alpha_i} + g_{t-i}^\delta \right) \\
\frac{\partial \lambda_t}{\partial \gamma_i} &= \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta \sum_{k=1}^p \alpha_k g_{t-k}^{\delta-1} (I_{t-k} + \gamma_k) \frac{\partial \lambda_{t-k}}{\partial \gamma_i} + \sum_{j=1}^q \beta_j \frac{\partial \lambda_{t-j}^\delta}{\partial \gamma_i} - \delta \alpha_i g_{t-i}^{\delta-1} (y_{t-i} - \lambda_{t-i}) \right), \\
\frac{\partial \lambda_t}{\partial \beta_j} &= \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta \sum_{i=1}^p \alpha_i g_{t-i}^{\delta-1} (I_{t-i} + \gamma_i) \frac{\partial \lambda_{t-i}}{\partial \beta_j} + \sum_{k=1}^q \beta_k \frac{\partial \lambda_{t-k}^\delta}{\partial \beta_j} + \lambda_{t-j}^\delta \right), \\
\frac{\partial \lambda_t}{\partial \delta} &= \frac{\lambda_t}{\delta \lambda_t^\delta} \left\{ \sum_{i=1}^p \alpha_i g_{t-i}^\delta \left( \frac{\delta}{g_{t-i}} (I_{t-i} + \gamma_i) \frac{\partial \lambda_{t-i}}{\partial \delta} + \ln(g_{t-i}) \right) + \sum_{j=1}^q \beta_j \frac{\partial \lambda_{t-j}^\delta}{\partial \delta} - \frac{\lambda_t^\delta}{\delta} \ln(\lambda_t^\delta) \right\},
\end{aligned}$$

where  $g_{t-i} = |y_{t-i} - \lambda_{t-i}| - \gamma_i(y_{t-i} - \lambda_{t-i})$  and  $I_t = \begin{cases} -1 & y_t > \lambda_t \\ 1 & y_t < \lambda_t \end{cases}$ . Thus, for the INAPARCH(1, 1) model the score function can then be explicitly written as

$$S_n(\theta) = \begin{bmatrix} \sum_{t=1}^n \left( \frac{y_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta_1} \\ \vdots \\ \sum_{t=1}^n \left( \frac{y_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta_5} \end{bmatrix}$$

with

$$\begin{aligned}
\frac{\partial \lambda_t}{\partial \theta_1} &\equiv \frac{\partial \lambda_t}{\partial \omega} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta (\alpha(I_{t-1} + \gamma) g_{t-1}^{\delta-1} + \beta \lambda_{t-1}^{\delta-1}) \frac{\partial \lambda_{t-1}}{\partial \omega} + 1 \right) \\
\frac{\partial \lambda_t}{\partial \theta_2} &\equiv \frac{\partial \lambda_t}{\partial \alpha} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta (\alpha(I_{t-1} + \gamma) g_{t-1}^{\delta-1} + \beta \lambda_{t-1}^{\delta-1}) \frac{\partial \lambda_{t-1}}{\partial \alpha} + g_{t-1}^\delta \right) \\
\frac{\partial \lambda_t}{\partial \theta_3} &\equiv \frac{\partial \lambda_t}{\partial \gamma} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta (\alpha(I_{t-1} + \gamma) g_{t-1}^{\delta-1} + \beta \lambda_{t-1}^{\delta-1}) \frac{\partial \lambda_{t-1}}{\partial \gamma} - \alpha \delta g_{t-1}^{\delta-1} (y_{t-1} - \lambda_{t-1}) \right) \\
\frac{\partial \lambda_t}{\partial \theta_4} &\equiv \frac{\partial \lambda_t}{\partial \beta} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta (\alpha(I_{t-1} + \gamma) g_{t-1}^{\delta-1} + \beta \lambda_{t-1}^{\delta-1}) \frac{\partial \lambda_{t-1}}{\partial \beta} + \lambda_{t-1}^\delta \right) \\
\frac{\partial \lambda_t}{\partial \theta_5} &\equiv \frac{\partial \lambda_t}{\partial \delta} = \frac{\lambda_t}{\delta \lambda_t^\delta} \left( \delta (\alpha(I_{t-1} + \gamma) g_{t-1}^{\delta-1} + \beta \lambda_{t-1}^{\delta-1}) \frac{\partial \lambda_{t-1}}{\partial \delta} + \alpha g_{t-1}^\delta \ln(g_{t-1}) + \beta \lambda_{t-1}^\delta \ln(\lambda_{t-1}) \right) - \frac{\lambda_t}{\delta} \ln(\lambda_t).
\end{aligned}$$

The solution of the equation  $S_n(\theta) = 0$  is the conditional maximum likelihood estimator,  $\hat{\theta}$ , if it exists. To study the asymptotic properties of the maximum likelihood estimator we proceed as follows: first it can be shown that the score function, evaluated at the true value of the parameter, say  $\theta = \theta_0$ , is asymptotically normal. The score function has martingale difference terms defined by

$$\frac{\partial \ell_t}{\partial \theta_i} = \left( \frac{y_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta_i}.$$

It follows that, at  $\theta = \theta_0$

$$E \left[ \frac{\partial \ell_t}{\partial \theta_0} | \mathcal{F}_{t-1} \right] = 0,$$

since  $E\left[\frac{y_t}{\lambda_t} - 1 | \mathcal{F}_{t-1}\right] = 0$ , and  $E\left[\left(\frac{y_t}{\lambda_t} - 1\right)^2 | \mathcal{F}_{t-1}\right] = V\left[\frac{y_t}{\lambda_t} - 1 | \mathcal{F}_{t-1}\right] = \frac{1}{\lambda_t}$ . It can also easily be shown that, for  $\delta \geq 2$

$$E\left[\lambda_t^{2-2\delta} | \mathcal{F}_{t-1}\right] < +\infty, \quad E\left[\lambda_t^{1-\delta} | \mathcal{F}_{t-1}\right] < +\infty$$

$$E\left[\lambda_t^{2-\delta} \ln(\lambda_t) | \mathcal{F}_{t-1}\right] < E[\ln(\lambda_t) | \mathcal{F}_{t-1}] < E[\lambda_t | \mathcal{F}_{t-1}] < +\infty$$

$$E\left[\lambda_t^2 \ln^2(\lambda_t) | \mathcal{F}_{t-1}\right] < +\infty, \quad E\left[\lambda_t \ln(\lambda_t) | \mathcal{F}_{t-1}\right] < +\infty.$$

Thus, it can be concluded that  $V\left[\frac{\partial \ell_t}{\partial \theta} | \mathcal{F}_{t-1}\right] < +\infty$  and that  $\partial \ell_t / \partial \theta$  is a martingale difference sequence with respect to  $\mathcal{F}_{t-1}$ . The application of a central limit theorem for martingales guarantees the desired asymptotic normality.

It is worth to mention here that in Section 2 it was concluded that the process has finite moments up to any positive order and is  $\tau$ -weak dependent, which implies ergodicity. This is sufficient to state that the Hessian matrix converges in probability to a finite limit. Finally, all third derivatives are bounded by a sequence that converges in probability. Given these three conditions, it is then concluded that the conditional maximum likelihood estimator,  $\hat{\theta}$ , is consistent and asymptotically normal,

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, G^{-1}(\theta_0)),$$

with variance-covariance matrix,  $G(\theta)$ , given by

$$G(\theta) = E\left[\frac{1}{\lambda_t} \left(\frac{\partial \lambda_t}{\partial \theta}\right) \left(\frac{\partial \lambda_t}{\partial \theta}\right)'\right].$$

A consistent estimator of  $G(\theta)$  is given by

$$G_n(\theta) = \sum_{t=1}^n V\left[\frac{\partial \ell_t(\theta)}{\partial \theta} | \mathcal{F}_{t-1}\right] = \sum_{t=1}^n \frac{1}{\lambda_t(\theta)} \left(\frac{\partial \lambda_t(\theta)}{\partial \theta}\right) \left(\frac{\partial \lambda_t(\theta)}{\partial \theta}\right)'.$$

## 4 Simulation

In this section, a simulation study computed using Matlab is carried out to illustrate the theoretical findings given above. This study contemplates five different combinations for  $\theta$ , which are displayed in Table 1 below. For each set of parameters, time series of length 500 with 300 independent replicates from the INAPARCH(1, 1) model were simulated.

Note that for C1-C4 cases, condition (5) holds, whereas for case C5 this condition fails. The results are summarized in Table 1 and the bias of the CML estimates are presented in Figure 2 for the combination of parameters C2 and C4. The point estimates for the  $\alpha$  parameter follow the theoretical values in a coherent way, even for very small theoretical values such as for the cases C1, C2 and C3. Nevertheless, the standard deviations of the point estimates are



relatively high, except for the C5 case.  $\gamma$  and  $\delta$  parameters are fairly estimated but there is a certain difficulty in the estimation of the  $\omega$  parameter, that tends to be underestimated, with the exception of the C5 case. Considering the  $\beta$  parameter, all point estimates are considerably higher than the theoretical values, the only exception being the C5 case, once again. There is also a very high degree of variability, in particular for  $\omega$  and  $\delta$  parameters.

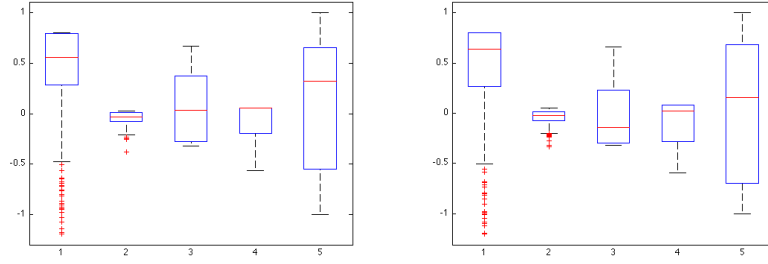
From this part of the simulation study a few conclusions can be drawn: firstly, it is clear that as the theoretical values of the parameters rise, the point estimates obtained are much closer to what was expected. Secondly, it seems that condition (5), the sufficient condition for ergodicity, does not seem to interfere with the quality of the point estimates for this model. In fact, best overall estimates were obtained for case C5, clearly not obeying condition (5).

**Table 1.** Parameter estimates and standard deviations (std) in parentheses.

Parameter Values							Point estimates and (std)				
Case	$\omega$	$\alpha$	$\gamma$	$\beta$	$\delta$	$2^{\delta}(2\alpha\delta + \frac{\beta}{2})$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\delta}$
C1	2.30	0.01	0.68	0.10	2.00	0.36	1.8510 (0.4825)	0.0641 (0.0685)	0.6356 (0.3180)	0.1850 (0.2246)	1.9245 (0.7156)
C2	2.30	0.03	0.68	0.06	2.00	0.60	1.9067 (0.5142)	0.0755 (0.0698)	0.6174 (0.3351)	0.1452 (0.1981)	1.9170 (0.6860)
C3	2.30	0.01	0.68	0.10	3.00	0.88	1.9674 (0.4229)	0.0571 (0.0684)	0.5922 (0.2914)	0.1572 (0.1813)	2.9588 (0.7183)
C4	2.30	0.05	0.68	0.08	2.00	0.96	1.8931 (0.5294)	0.0880 (0.0722)	0.7005 (0.3070)	0.1753 (0.2102)	1.9535 (0.7154)
C5	2.30	0.30	0.68	0.10	2.00	5.00	2.2724 (0.7519)	0.3082 (0.1290)	0.7489 (0.2229)	0.1294 (0.1318)	2.0401 (0.6510)

#### 4.1 Log-likelihood analysis

For C2 and C4 cases, 300 samples were simulated considering values of  $\delta$  varying from 2.0 to 3.0 (i.e., six different situations for each case). After preliminary data analysis with the construction of boxplots and histograms that confirm the presence of overdispersion, the log-likelihood was studied in the following manner: for each set of 300 samples the log-likelihood was calculated, varying the  $\delta$  parameter in the range 2.0 to 3.0. It was expected that the log-likelihood was maximum for the  $\delta$  value used to simulate that particular set of 300 samples. Results are presented in Table 2 for Case 2. Case 2 was chosen for representation herein just because for this case the first three values for the  $\delta$  parameter lie inside the region that obeys condition (5) and the last 3 lie outside this region. Nevertheless, same behaviour was observed for both Case 2 and Case 4 and the  $\delta$  value for which the calculated log-likelihood was maximum was exactly what



**Fig. 2.** Bias of the conditional ML estimates, for the combination of parameters C2 (*left*) and C4 (*right*). Numbers one to five below the boxplots refer to the estimated parameters, in the order appearing in Table 1.

was expected for both cases and all 6 different situations. In Table 2, it can be observed that the mean log-likelihood is maximum for the  $\delta$  value corresponding to the  $\delta$  value used for the simulation of the respective set of samples.

**Table 2.** Maximum likelihood estimation results for Case 2.

Samples simulated with	Log-likelihood for varying $\delta$					
$\theta=(2.30, 0.03, 0.68, 0.06, \delta)$	2.0	2.2	2.4	2.6	2.8	3.0
$\delta = 2.00$	-785.4787	-786.1563	-787.6991	-789.6634	-791.8038	-793.9828
$\delta = 2.20$	-775.2089	-774.5939	-775.0658	-776.1291	-777.5016	-779.0191
$\delta = 2.40$	-766.7914	-765.1027	-764.6847	-764.9993	-765.7337	-766.7013
$\delta = 2.60$	-760.1167	-757.5743	-756.4490	-756.1685	-756.3958	-756.9265
$\delta = 2.80$	-755.0275	-751.7783	-750.0676	-749.2947	-749.1024	-749.2715
$\delta = 3.00$	-751.1783	-747.3026	-745.0736	-743.8653	-743.3025	-743.1530

## 5 Real-Data Example: Transaction Modeling

In this section, the results above are applied in the analysis of the motivating examples presented in Figure 1, Section 1. As already described, data consist on the number of transactions per minute during one trading day for Glaxosmithkline and Astrazeneca. CML estimation method was applied and the results are shown in Table 3. Note that the estimated value of  $\gamma$  is negative for both time series meaning that there is evidence that positive shocks have stronger impact on overdispersion than negative shocks. Another important feature exhibited by both time series is that the estimated value of  $\delta$  fail the condition  $\delta \geq 2$ . It is

**Table 3.** Maximum likelihood estimation results for Glaxosmithkline and Astrazeneca time series. Standard errors in parentheses.

Time series	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\delta}$
Glaxosmithkline	0.3781 (0.0685)	0.1392 (0.0074)	-0.3269 (0.0843)	0.8791 (0.0073)	0.9826 (0.0005)
Astrazeneca	2.4862 (0.1087)	0.2824 (0.0062)	-0.2787 (0.0363)	0.7501 (0.0044)	1.0598 (0.0008)

worth mentioning that this is not a surprising result since in the estimation of the Standard & Poor 500 stock market daily closing price index in [4], the  $\delta$  estimate obtained did not also satisfy such sufficient condition for the process to be covariance stationary.

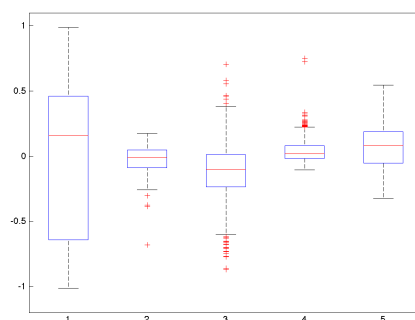
A short simulation study was also carried out in this section. The CML point estimates of both real-data series in Table 3 were used to simulate 300 independent replicates of length 500 from the INAPARCH(1,1) model, namely, GSK and AZN cases, referring, respectively, to the samples based on the point estimates for the Glaxosmithkline and Astrazeneca time series. CML estimates were then obtained for these samples and the results are presented in Table 4, with corresponding bias in Figure 3. Regarding Figure 3, it can be seen that

**Table 4.** Maximum likelihood estimation results for GSK and AZN cases. Standard deviations in parentheses.

Samples	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\delta}$
GSK Case	0.7396 (0.5779)	0.1820 (0.0940)	-0.2417 (0.2596)	0.8710 (0.0257)	1.2202 (0.3882)
AZN Case	2.4955 (0.6659)	0.3094 (0.1019)	-0.1473 (0.2586)	0.6996 (0.1074)	0.9987 (0.1612)

variability and the tendency to underestimate the  $\omega$  parameter is maintained (taking in consideration the median value) but in relation to the  $\delta$  parameter variability has decreased significantly. From inspection of Table 4, it can be said that, in general, CML point estimates are not very far from what was expected in both cases, although better overall estimates were obtained for the AZN case. Considering that condition (5) was not fulfilled for either AZN or GSK cases ( $2^\delta(2\alpha\delta + \frac{\beta}{2})$  equals 2.0297 for the AZN case and 1.4091 for the GSK case), as was already mentioned in Section 4, it seems that violating the sufficient condition for ergodicity has no effect on the behaviour of the estimation procedure. The impact of violating necessary instead of sufficient conditions for ergodicity remains as a topic of future work.

**Acknowledgments.** This research was partially supported by Portuguese funds through the Center for Research and Development in Mathematics and Appli-



**Fig. 3.** Bias of CML estimates, for AZN Case. Numbers one to five below the boxplots refer to the estimated parameters, in the order appearing in Table 4.

cations, CIDMA, and the Portuguese Foundation for Science and Technology, FCT–Fundação para a Ciência e a Tecnologia, within project UID/MAT/04106/2013.

## References

1. Brännäs, K., Quoreshi, A.M.M.S.: Integer-valued moving average modelling of the number of transactions in stocks. *Appl. Financ. Econ.* 20(18), 1429–1440 (2010)
2. Davis, R.A., Dunsmuir, W.T.M., Streett, S.B.: Observation driven models for poisson counts. *Biomet.* 90, 777–790 (2003)
3. Dedecker, J., Prieur, C.: Coupling for  $\tau$ -dependent sequences and applications. *J. Theor. Probab.* 17, 861–885 (2004)
4. Ding, Z., Granger, C.W., Engle, R.F.: A long memory property of stock market returns and a new model. *J. Empir. Financ.* 1(1), 83–106 (1993)
5. Doukhan, P., Fokianos, K., Tjøstheim, D.: On weak dependence conditions for Poisson autoregressions. *Statist. Probab. Lett.* 82, 942–948 (2012)
6. Ferland, R., Latour, A., Oraichi, D.: Integer-valued GARCH process. *J. Time Ser. Anal.* 6, 923–942 (2006)
7. Fokianos, K., Tjøstheim, D.: Nonlinear Poisson autoregression. *Ann. Inst. Statist. Math.* 64, 1205–1225 (2012)
8. Fokianos, K., Rahbek, A., Tjøstheim, D.: Poisson autoregression. *J. Am. Stat. Assoc.* 104, 1430–1439 (2009)
9. Franke, J.: Weak dependence of functional INGARCH processes. Technical Report 126, Technische Universität Kaiserslautern (2010)
10. Heinen, A.: Modelling time series count data: an autoregressive conditional poisson model. Center for Operations Research and Econometrics (CORE) Discussion Paper No. 2003-63, University of Louvain, Belgium (2003)
11. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer-Verlag, New York (1994)
12. Neumann, M.H.: Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* 17, 1268–1284 (2011)